

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Discrete-Time Smoothing

7.1. Introduction

Observations are invariably accompanied by measurement noise and optimal filters are the usual solution of choice. Filter performances that fall short of user expectations motivate the pursuit of smoother solutions. Smoothers promise useful mean-square-error improvement at mid-range signal-to-noise ratios, provided that the assumed model parameters and noise statistics are correct.

In general, discrete-time filters and smoothers are more practical than the continuous-time counterparts. Often a designer may be able to value-add by assuming low-order discrete-time models which bear little or no resemblance to the underlying processes. Continuous-time approaches may be warranted only when application-specific performance considerations outweigh the higher overheads.

This chapter canvasses the main discrete-time fixed-point, fixed-lag and fixed interval smoothing results [1] – [9]. Fixed-point smoothers [1] calculate an improved estimate at a prescribed past instant in time. Fixed-lag smoothers [2] – [3] find application where small end-to-end delays are tolerable, for example, in press-to-talk communications or receiving public broadcasts. Fixed-interval smoothers [4] – [9] dispense with the need to fine tune the time of interest or the smoothing lags. They are suited to applications where processes are staggered such as delayed control or off-line data analysis. For example, in underground coal mining, smoothed position estimates and control signals can be calculated while a longwall shearer is momentarily stationary at each end of the face [9]. Similarly, in exploration drilling, analyses are typically carried out post-data acquisition.

The smoother descriptions are organised as follows. Section 7.2 sets out two prerequisites: time-varying adjoint systems and Riccati difference equation comparison theorems. Fixed-point, fixed-lag and fixed-interval smoothers are discussed in Sections 7.3, 7.4 and 7.5, respectively. It turns out that the structures of the discrete-time smoothers are essentially the same as those of the previously-described continuous-time versions. Differences arise in the calculation of Riccati equation solutions and the gain matrices. Consequently, the treatment

“An inventor is simply a person who doesn't take his education too seriously. You see, from the time a person is six years old until he graduates from college he has to take three or four examinations a year. If he flunks once, he is out. But an inventor is almost always failing. He tries and fails maybe a thousand times. If he succeeds once then he's in. These two things are diametrically opposite. We often say that the biggest job we have is to teach a newly hired employee how to fail intelligently. We have to train him to experiment over and over and to keep on trying and failing until he learns what will work.”
Charles Franklin Kettering

is somewhat condensed. It is reaffirmed that the above-mentioned smoothers outperform the Kalman filter and the minimum-variance smoother provides the best performance.

7.2. Prerequisites

7.2.1 Time-varying Adjoint Systems

Consider a linear time-varying system, \mathcal{G} , operating on an input, w , namely, $y = \mathcal{G} w$. Here, w denotes the set of w_k over an interval $k \in [0, N]$. It is assumed that $\mathcal{G} : \mathbb{R}^p \rightarrow \mathbb{R}^q$ has the state-space realisation

$$x_{k+1} = A_k x_k + B_k w_k, \quad (1)$$

$$y_k = C_k x_k + D_k w_k. \quad (2)$$

As before, the adjoint system, \mathcal{G}^H , satisfies

$$\langle y, \mathcal{G} w \rangle = \langle \mathcal{G}^H y, w \rangle \quad (3)$$

for all $y \in \mathbb{R}^q$ and $w \in \mathbb{R}^p$.

Lemma 1: In respect of the system \mathcal{G} described by (1) – (2), with $x_0 = 0$, the adjoint system \mathcal{G}^H having the realisation

$$\zeta_{k-1} = A_k^T \zeta_k - C_k^T u_k, \quad (4)$$

$$z_k = -B_k^T \zeta_k + D_k^T u_k, \quad (5)$$

with $\zeta_N = 0$, satisfies (3).

A proof appears in [7] and proceeds similarly to that within Lemma 1 of Chapter 2. The simplification $D_k = 0$ is assumed below unless stated otherwise.

7.2.2 Riccati Equation Comparison

The ensuing performance comparisons of filters and smoothers require methods for comparing the solutions of Riccati difference equations which are developed below. Simplified Riccati difference equations which do not involve the B_k and measurement noise covariance matrices are considered initially. A change of variables for the more general case is stated subsequently.

Suppose there exist $A_{t+k-1} \in \mathbb{R}^{n \times n}$, $\tilde{C}_{t+k-1} \in \mathbb{R}^{p \times n}$, $\tilde{Q}_{t+k-1} = \tilde{Q}_{t+k-1}^T \in \mathbb{R}^{n \times n}$ and $P_{t+k-1} = P_{t+k-1}^T \in \mathbb{R}^{n \times n}$ for a $t \geq 0$ and $k \geq 0$. Following the approach of Wimmer [10], define the Riccati operator

$$\begin{aligned} \Phi(P_{t+k-1}, A_{t+k-1}, \tilde{C}_{t+k-1}, \tilde{Q}_{t+k-1}) &= A_{t+k-1} P_{t+k-1} A_{t+k-1}^T + \tilde{Q}_{t+k-1} \\ &\quad - A_{t+k-1} P_{t+k-1} \tilde{C}_{t+k-1}^T (I + \tilde{C}_{t+k-1} P_{t+k-1} \tilde{C}_{t+k-1}^T)^{-1} \tilde{C}_{t+k-1} P_{t+k-1} A_{t+k-1}^T. \end{aligned} \quad (6)$$

“If you’re not failing every now and again, it’s a sign you’re not doing anything very innovative.”
(Woody) Allen Stewart Konigsberg

Let $\Gamma_{t+k-1} = \begin{bmatrix} A_{t+k-1} & -\tilde{C}_{t+k-1}^T \tilde{C}_{t+k-1} \\ -\tilde{Q}_{t+k-1} & -A_{t+k-1}^T \end{bmatrix}$ denote the Hamiltonian matrix corresponding to $\Phi(P_{t+k-1}, A_{t+k-1}, \tilde{C}_{t+k-1}, \tilde{Q}_{t+k-1})$ and define $J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$, in which I is an identity matrix of appropriate dimensions. It is known that solutions of (6) are monotonically dependent on $J\Gamma_{k+1} = \begin{bmatrix} \tilde{Q}_{t+k-1} & A_{t+k-1}^T \\ A_{t+k-1} & -\tilde{C}_{t+k-1}^T \tilde{C}_{t+k-1} \end{bmatrix}$. Consider a second Riccati operator employing the same initial solution P_{t+k-1} but different state-space parameters

$$\begin{aligned} \Phi(P_{t+k-1}, A_{t+k}, \tilde{C}_{t+k}, \tilde{Q}_{t+k}) &= A_{t+k} P_{t+k-1} A_{t+k}^T + \tilde{Q}_{t+k} \\ &\quad - A_{t+k} P_{t+k-1} \tilde{C}_{t+k}^T (I + \tilde{C}_{t+k} P_{t+k-1} \tilde{C}_{t+k}^T)^{-1} \tilde{C}_{t+k} P_{t+k-1} A_{t+k}^T. \end{aligned} \quad (7)$$

The following theorem, which is due to Wimmer [10], compares the above two Riccati operators.

Theorem 1: [10]: Suppose that

$$\begin{bmatrix} \tilde{Q}_{t+k-1} & A_{t+k-1}^T \\ A_{t+k-1} & -\tilde{C}_{t+k-1}^T \tilde{C}_{t+k-1} \end{bmatrix} \geq \begin{bmatrix} \tilde{Q}_{t+k} & A_{t+k}^T \\ A_{t+k} & -\tilde{C}_{t+k}^T \tilde{C}_{t+k} \end{bmatrix}$$

for a $t \geq 0$ and for all $k \geq 0$. Then

$$\Phi(P_{t+k-1}, A_{t+k-1}, \tilde{C}_{t+k-1}, \tilde{Q}_{t+k-1}) \geq \Phi(P_{t+k-1}, A_{t+k}, \tilde{C}_{t+k}, \tilde{Q}_{t+k}) \quad (8)$$

for all $k \geq 0$.

The above result underpins the following more general Riccati difference equation comparison theorem.

Theorem 2: [11], [8]: With the above definitions, suppose for a $t \geq 0$ and for all $k \geq 0$ that:

- (i) there exists a $P_t \geq P_{t+1}$ and
- (ii) $\begin{bmatrix} \tilde{Q}_{t+k-1} & A_{t+k-1}^T \\ A_{t+k-1} & -\tilde{C}_{t+k-1}^T \tilde{C}_{t+k-1} \end{bmatrix} \geq \begin{bmatrix} \tilde{Q}_{t+k} & A_{t+k}^T \\ A_{t+k} & -\tilde{C}_{t+k}^T \tilde{C}_{t+k} \end{bmatrix}$.

Then $P_{t+k} \geq P_{t+k+1}$ for all $k \geq 0$.

Proof: Assumption (i) is the $k = 0$ case for an induction argument. For the inductive step, denote $P_{t+k} = \Phi(P_{t+k-1}, A_{t+k-1}, \tilde{C}_{t+k-1}, \tilde{Q}_{t+k-1})$ and $P_{t+k+1} = \Phi(P_{t+k}, A_{t+k}, \tilde{C}_{t+k}, \tilde{Q}_{t+k})$. Then

"Although personally I am quite content with existing explosives, I feel we must not stand in the path of improvement." Winston Leonard Spencer-Churchill

$$P_{t+k} - P_{t+k+1} = (\Phi(P_{t+k-1}, A_{t+k-1}, \tilde{C}_{t+k-1}, \tilde{Q}_{t+k-1}) - \Phi(P_{t+k-1}, A_{t+k}, \tilde{C}_{t+k}, \tilde{Q}_{t+k})) \quad (9)$$

$$+ (\Phi(P_{t+k-1}, A_{t+k}, \tilde{C}_{t+k}, \tilde{Q}_{t+k}) - \Phi(P_{t+k}, A_{t+k}, \tilde{C}_{t+k}, \tilde{Q}_{t+k}))$$

The first term on the right-hand-side of (9) is non-negative by virtue of Assumption (ii) and Theorem 1. By appealing to Theorem 2 of Chapter 5, the second term on the right-hand-side of (9) is non-negative and thus $P_{t+k} - P_{t+k+1} \geq 0$. \square

A change of variables [8] $\tilde{C}_k = R_k^{-1/2} C_k$ and $\tilde{Q}_k = B_k Q_k B_k^T$, allows the application of Theorem 2 to the more general forms of Riccati differential equations.

7.3 Fixed-Point Smoothing

7.3.1 Solution Derivation

The development of a discrete-time fixed-point smoother follows the continuous-time case. An innovation by Zachrisson [12] involves transforming the smoothing problem into a filtering problem that possesses an augmented state. Following the approach in [1], consider

an augmented state vector $x_k^{(a)} = \begin{bmatrix} x_k \\ \xi_k \end{bmatrix}$ for the signal model

$$x_{k+1}^{(a)} = A_k^{(a)} x_k^{(a)} + B_k^{(a)} w_k, \quad (10)$$

$$z_k = C_k^{(a)} x_k^{(a)} + v_k, \quad (11)$$

where $A_k^{(a)} = \begin{bmatrix} A_k & 0 \\ 0 & I \end{bmatrix}$, $B_k^{(a)} = \begin{bmatrix} B_k \\ 0 \end{bmatrix}$ and $C_k^{(a)} = [C_k \ 0]$. It can be seen that the first component of $x_k^{(a)}$ is x_k , the state of the system $x_{k+1} = A_k x_k + B_k w_k$, $y_k = C_k x_k + v_k$. The second component, ξ_k , equals x_k at time $k = \tau$, that is, $\xi_k = x_\tau$. The objective is to calculate an estimate $\hat{\xi}_k$ of ξ_k at time $k = \tau$ from measurements z_k over $k \in [0, N]$. A solution that minimises the variance of the estimation error is obtained by employing the standard Kalman filter recursions for the signal model (10) – (11). The predicted and corrected states are respectively obtained from

$$\hat{x}_{k/k}^{(a)} = (I - L_k^{(a)} C_k^{(a)}) \hat{x}_{k/k-1}^{(a)} + L_k^{(a)} z_k, \quad (12)$$

$$\hat{x}_{k+1/k}^{(a)} = A_k^{(a)} \hat{x}_{k/k}^{(a)} \quad (13)$$

$$= (A_k^{(a)} - K_k^{(a)} C_k^{(a)}) \hat{x}_{k/k-1}^{(a)} + K_k^{(a)} z_k, \quad (14)$$

where $K_k = A_k^{(a)} L_k^{(a)}$ is the predictor gain, $L_k^{(a)} = P_{k/k-1}^{(a)} (C_k^{(a)})^T (C_k^{(a)} P_{k/k-1}^{(a)} (C_k^{(a)})^T + R_k)^{-1}$ is the filter gain,

“Never before in history has innovation offered promise of so much to so many in so short a time.”
William Henry (Bill) Gates III

$$P_{k/k}^{(a)} = P_{k/k-1}^{(a)} - P_{k/k-1}^{(a)} (C_k^{(a)})^T (L_k^{(a)})^T \quad (15)$$

is the corrected error covariance and

$$P_{k+1/k}^{(a)} = A_k^{(a)} P_{k/k}^{(a)} (A_k^{(a)})^T + B_k^{(a)} Q_k (B_k^{(a)})^T \quad (16)$$

$$\begin{aligned} &= A_k^{(a)} P_{k/k-1}^{(a)} (A_k^{(a)})^T - A_k^{(a)} P_{k/k-1}^{(a)} (C_k^{(a)})^T (K_k^{(a)})^T + B_k^{(a)} Q_k (B_k^{(a)})^T \\ &= A_k^{(a)} P_{k/k-1}^{(a)} \left((A_k^{(a)})^T - (C_k^{(a)})^T (K_k^{(a)})^T \right) + B_k^{(a)} Q_k (B_k^{(a)})^T \end{aligned} \quad (17)$$

is the predicted error covariance. The above Riccati difference equation is written in the partitioned form

$$\begin{aligned} P_{k+1/k}^{(a)} &= \begin{bmatrix} P_{k+1/k} & \Sigma_{k+1/k}^T \\ \Sigma_{k+1/k} & \Omega_{k+1/k} \end{bmatrix} \\ &= \begin{bmatrix} A_k & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_{k/k-1} & \Sigma_{k/k-1}^T \\ \Sigma_{k/k-1} & \Omega_{k/k-1} \end{bmatrix} \\ &\quad \times \left(\begin{bmatrix} A_k^T & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} C_k^T \\ 0 \end{bmatrix} \begin{bmatrix} K_k^T & \underline{K}_k^T \end{bmatrix} (C_k P_{k/k-1} C_k^T + R_k)^{-1} \right) + \begin{bmatrix} B_k \\ 0 \end{bmatrix} Q_k \begin{bmatrix} B_k^T & 0 \end{bmatrix}, \end{aligned} \quad (18)$$

in which the gains are given by

$$\begin{aligned} K_k^{(a)} &= \begin{bmatrix} K_k \\ \underline{L}_k \end{bmatrix} = \begin{bmatrix} A_k & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_{k/k-1} & \Sigma_{k/k-1}^T \\ \Sigma_{k/k-1} & \Pi_{k/k-1} \end{bmatrix} \begin{bmatrix} C_k^T \\ 0 \end{bmatrix} (C_k P_{k/k-1} C_k^T + R_k)^{-1} \\ &= \begin{bmatrix} A_k P_{k/k-1} C_k^T \\ \Sigma_{k/k-1} C_k^T \end{bmatrix} (C_k P_{k/k-1} C_k^T + R_k)^{-1}, \end{aligned} \quad (19)$$

see also [1]. The predicted error covariance components can be found from (18), viz.,

$$P_{k+1/k} = A_k P_{k/k} A_k^T + B_k Q_k B_k^T, \quad (20)$$

$$\Sigma_{k+1/k} = \Sigma_{k/k-1} (A_k^T - C_k^T K_k^T), \quad (21)$$

$$\Omega_{k+1/k} = \Omega_{k/k-1} - \Sigma_{k/k-1} C_k^T \underline{L}_k^T. \quad (22)$$

The sequences (21) – (22) can be initialised with $\Sigma_{\tau+1/\tau} = P_{\tau/\tau}$ and $\Omega_{\tau+1/\tau} = P_{\tau/\tau}$. The state corrections are obtained from (12), namely,

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k (z_k - C_k \hat{x}_{k/k-1}), \quad (23)$$

$$\hat{\xi}_{k/k} = \hat{\xi}_{k/k-1} + \underline{L}_k (z_k - C_k \hat{x}_{k/k-1}). \quad (24)$$

Similarly, the state predictions follow from (13),

“You can’t just ask customers what they want and then try to give that to them. By the time you get it built, they’ll want something new.” *Steven Paul Jobs*

$$\hat{x}_{k+1/k} = A_k \hat{x}_{k/k}, \quad (25)$$

$$\hat{\xi}_{k+1/k} = \hat{\xi}_{k/k}. \quad (26)$$

In summary, the fixed-point smoother estimates for $k \geq \tau$ are given by (24), which is initialised by $\hat{\xi}_{\tau/\tau} = \hat{x}_{\tau/\tau}$. The smoother gain is calculated as $\underline{L}_k = \Sigma_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + R_k)^{-1}$, where $\Sigma_{k/k-1}$ is given by (21).

7.3.2 Performance

It follows from the above that $\Omega_{k+1/k} = \Omega_{k/k}$ and so

$$\Omega_{k+1/k+1} = \Omega_{k/k} - \Sigma_{k/k-1} C_k^T \underline{L}_k^T. \quad (27)$$

Next, it is argued that the discrete-time fixed-point smoother provides a performance improvement over the filter.

Lemma 2 [1]: *In respect of the fixed point smoother (24),*

$$P_{\tau/\tau} \geq \Omega_{k/k}. \quad (28)$$

Proof: *The recursion (22) may be written as the sum*

$$\Omega_{k+1/k+1} = \Omega_{\tau/\tau} - \sum_{i=\tau}^k \Sigma_{i/i-1} C_i^T (C_i P_{i/i-1} C_i^T + R)^{-1} C_i \Sigma_{i/i-1}, \quad (29)$$

where $\Omega_{\tau/\tau} = P_{\tau/\tau}$. Hence, $P_{\tau/\tau} - \Omega_{k+1/k+1} = \sum_{i=\tau}^k \Sigma_{i/i-1} C_i^T (C_i P_{i/i-1} C_i^T + R)^{-1} C_i \Sigma_{i/i-1} \geq 0$. \square

Example 1. Consider a first-order time-invariant plant, in which $A = 0.9$, $B = 1$, $C = 0.1$ and $Q = 1$. An understanding of a fixed-point smoother's performance can be gleaned by examining the plots of the $\Sigma_{k/k}$ and $\Omega_{k/k}$ sequences shown in Fig. 1(a) and (b), respectively. The bottom lines of the figures correspond to measurement noise covariances of $R = 0.01$ and the top lines correspond to $R = 5$. It can be seen for this example, that the $\Sigma_{k/k}$ have diminishing impact after about 15 samples beyond the point of interest. From Fig. 1(b), it can be seen that smoothing appears most beneficial at mid-range measurement noise power, such as $R = 0.2$, since the plots of $\Omega_{k/k}$ become flatter for $R \geq 1$ and $R \leq 0.05$.

"There are no big problems, there are just a lot of little problems." Henry Ford

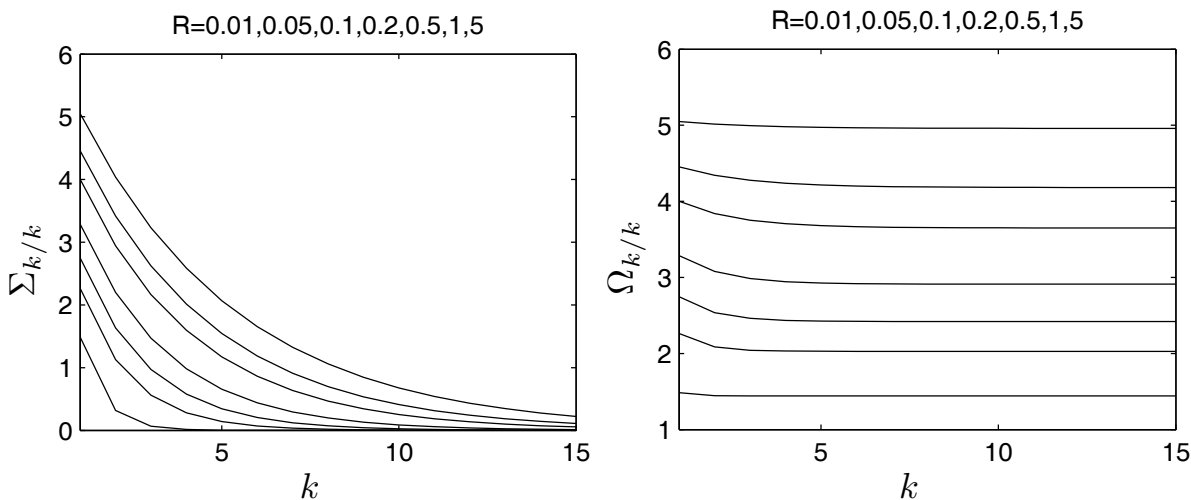


Figure 1(a). Smoother estimation variances $\Sigma_{k/k}$ versus k for Example 1.

Figure 1(b). Smoother estimation variances $\Omega_{k+1/k+1}$ versus k for Example 1.

7.4 Fixed-Lag Smoothing

7.4.1 High-order Solution

Discrete-time fixed-lag smoothers calculate state estimates, $\hat{x}_{k-N/k}$, at time k given a delay of N steps. The objective is to minimise $E\{(x_{k-N} - \hat{x}_{k-N/k})(x_{k/N-1} - \hat{x}_{k-N/k})^T\}$. A common solution approach is to construct an augmented signal model that includes delayed states and then apply the standard Kalman filter recursions, see [1] – [3] and the references therein. Consider the signal model

$$\begin{bmatrix} x_{k+1} \\ x_k \\ x_{k-1} \\ \vdots \\ x_{k+1-N} \end{bmatrix} = \begin{bmatrix} A_k & 0 & \cdots & 0 \\ I_N & & & 0 \\ 0 & I_N & & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & & I_N & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \\ x_{k-2} \\ \vdots \\ x_{k-N} \end{bmatrix} + \begin{bmatrix} B_k \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{30}$$

and

$$z_k = \begin{bmatrix} C_k & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \\ x_{k-2} \\ \vdots \\ x_{k-N} \end{bmatrix} + v_k. \tag{31}$$

“If the only tool you have is a hammer, you tend to see every problem as a nail.” *Abraham Maslow*

By applying the Kalman filter recursions to the above signal model, the predicted states are obtained as

$$\begin{bmatrix} \hat{x}_{k+1} \\ \hat{x}_k \\ \hat{x}_{k-1} \\ \vdots \\ \hat{x}_{k+1-N} \end{bmatrix} = \begin{bmatrix} A_k & 0 & \cdots & 0 \\ I_N & & & 0 \\ 0 & I_N & & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & I_N & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_k \\ \hat{x}_{k-1} \\ \hat{x}_{k-2} \\ \vdots \\ \hat{x}_{k-N} \end{bmatrix} + \begin{bmatrix} K_{0,k} \\ K_{1,k} \\ K_{2,k} \\ \vdots \\ K_{N,k} \end{bmatrix} (z_k - C_k \hat{x}_{k/k-1}), \quad (32)$$

where $K_{0,k}$, $K_{1,k}$, $K_{2,k}$, ..., $K_{N,k}$ denote the submatrices of the predictor gain. Two important observations follow from the above equation. First, the desired smoothed estimates $\hat{x}_{k-1/k}$... $\hat{x}_{k-N+1/k}$ are contained within the one-step-ahead prediction (32). Second, the fixed lag-smoother (32) inherits the stability properties of the original Kalman filter.

7.4.2 Reduced-order Solution

Equation (32) is termed a high order solution because the dimension of the above augmented state matrix is $(N+2)n \times (N+2)n$. Moore [1] - [3] simplified (32) to obtain elegant reduced order solution structures as follows. Let

$$\begin{bmatrix} P_{k+1/k}^{(0,0)} & P_{k+1/k}^{(0,1)} & \cdots & P_{k+1/k}^{(0,N)} \\ P_{k+1/k}^{(1,0)} & P_{k+1/k}^{(1,1)} & & \vdots \\ \vdots & & \ddots & \\ P_{k+1/k}^{(N,0)} & \cdots & & P_{k+1/k}^{(N,N)} \end{bmatrix}, \quad P_{k+1/k}^{(i,j)} = (P_{k+1/k}^{(j,i)})^T,$$

denote the predicted error covariance matrix. For $0 \leq i \leq N$, the smoothed states within (32) are given by

$$\hat{x}_{k-i/k} = \hat{x}_{k-i/k-1} + K_{i+1,k} (z_k - C_k \hat{x}_{k/k-1}), \quad (33)$$

where

$$K_{i+1,k} = P_{k+1/k}^{(i,0)} C_k^T (C_k P_{k+1/k}^{(0,0)} C_k^T + R_k)^{-1}. \quad (34)$$

Recursions for the error covariance submatrices of interest are

$$P_{k+1/k}^{(i+1,0)} = P_{k+1/k}^{(i,0)} (A_k - K_{0,k} C_k)^T, \quad (35)$$

$$P_{k+1/k}^{(i+1,i+1)} = P_{k+1/k}^{(i,i)} - P_{k+1/k}^{(i,0)} C_k K_{i+1,k}^T. \quad (36)$$

Another rearrangement of (33) - (34) to reduce the calculation cost further is described in [1].

"You have to seek the simplest implementation of a problem solution in order to know when you've reached your limit in that regard. Then it's easy to make tradeoffs, to back off a little, for performance reasons." *Stephen Gary Wozniak*

7.4.3 Performance

Two facts that stem from (36) are stated below.

Lemma 3: In respect of the fixed-lag smoother (33) – (36), the following applies.

- (i) The error-performance improves with increasing smoothing lag.
- (ii) The fixed-lag smoothers outperform the Kalman filter.

Proof:

- (i) The claim follows by inspection of (34) and (36).
- (ii) The observation follows by recognising that $P_{k+1/k}^{(1,1)} = E\{(x_k - \hat{x}_{k/k})(x_k - \hat{x}_{k/k})^T\}$ within (i). \square

It can also be seen from the term $-P_{k+1/k}^{(i,0)}C_k^T(C_kP_{k+1/k}^{(0,0)}C_k^T + R_k)^{-1}C_kP_{k+1/k}^{(i,0)}$ within (36) that the benefit of smoothing diminishes as R_k becomes large.

7.5 Fixed-Interval Smoothing

7.5.1 The Maximum-Likelihood Smoother

7.5.1.1 Solution Derivation

The most commonly used fixed-interval smoother is undoubtedly the solution reported by Rauch [5] in 1963 and two years later with Tung and Striebel [6]. Although this smoother does not minimise the error variance, it has two desirable attributes. First, it is a low-complexity state estimator. Second, it can provide close to optimal performance whenever the accompanying assumptions are reasonable.

The smoother involves two passes. In the first (forward) pass, filtered state estimates, $\hat{x}_{k/k}$, are calculated from

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - C_k\hat{x}_{k/k-1}), \quad (37)$$

$$\hat{x}_{k+1/k} = A_k\hat{x}_{k/k}, \quad (38)$$

where $L_k = P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}$ is the filter gain, $K_k = A_kL_k$ is the predictor gain, in which $P_{k/k} = P_{k/k-1} - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1}$ and $P_{k+1/k} = A_kP_{k/k}A_k^T + B_kQ_kB_k^T$. In the second backward pass, Rauch, Tung and Striebel calculate smoothed state estimates, $\hat{x}_{k/N}$, from the beautiful one-line recursion

$$\hat{x}_{k/N} = \hat{x}_{k/k} + G_k(\hat{x}_{k+1/N} - \hat{x}_{k+1/k}), \quad (39)$$

where

$$G_k = P_{k/k}A_{k+1}^T P_{k+1/k}^{-1} \quad (40)$$

is the smoother gain. The above sequence is initialised by $\hat{x}_{k/N} = \hat{x}_{k/k}$ at $k = N$. In the first public domain appearance of (39), Rauch [5] referred to a Lockheed Missile and Space

“For every problem there is a solution which is simple, clean and wrong.” Henry Louis Mencken

Company Technical Report co-authored with Tung and Striebel. Consequently, (39) is commonly known as the Rauch-Tung-Striebel smoother. This smoother was derived in [6] using the maximum-likelihood method and an outline is provided below.

The notation $x_k \sim \mathcal{N}(\mu, R_{xx})$ means that a discrete-time random variable x_k with mean μ and covariance R_{xx} has the normal (or Gaussian) probability density function

$$p(x_k) = \frac{1}{(2\pi)^{n/2} |R_{xx}|^{1/2}} \exp\left\{-0.5(x_k - \mu)^T R_{xx}^{-1}(x_k - \mu)\right\}. \quad (41)$$

Rauch, Tung and Striebel assumed that [6]:

$$\hat{x}_{k+1/N} \sim \mathcal{N}(A_k \hat{x}_{k/N}, B_k Q_k B_k^T). \quad (42)$$

$$\hat{x}_{k/N} \sim \mathcal{N}(\hat{x}_{k/k}, P_{k/k}). \quad (43)$$

From the approach of [6], setting the partial derivative of the logarithm of the joint density function to zero results in

$$0 = \frac{\partial(\hat{x}_{k+1/N} - A_k \hat{x}_{k/N})^T}{\partial \hat{x}_{k/N}} (B_k Q_k B_k^T)^{-1} (\hat{x}_{k+1/N} - A_k \hat{x}_{k/N})^T \frac{n!}{r!(n-r)!} + \frac{\partial(\hat{x}_{k/N} - \hat{x}_{k/N})^T}{\partial \hat{x}_{k/N}} P_{k/k}^{-1} (\hat{x}_{k/N} - \hat{x}_{k/N}).$$

Rearranging the above equation leads to

$$\hat{x}_{k/N} = (I + P_{k/k} A_k^T (B_k Q_k B_k^T)^{-1} A_k)^{-1} (\hat{x}_{k/N} + P_{k/k} A_k^T B_k Q_k B_k^T)^{-1} \hat{x}_{k/N}^T. \quad (44)$$

From the Matrix Inversion Lemma

$$(I + P_{k/k} A_k^T (B_k Q_k B_k^T)^{-1} A_k)^{-1} = I - G_k A_k. \quad (45)$$

The solution (39) is found by substituting (45) into (44). Some further details of Rauch, Tung and Striebel's derivation appear in [13].

7.5.1.2 Alternative Forms

The smoother gain (40) can be calculated in different ways. Assuming that A_k is non-singular, it follows from $P_{k+1/k} = A_k P_{k/k} A_k^T + B_k Q_k B_k^T$ that $P_{k/k} A_k^T P_{k+1/k}^{-1} = A_k^{-1} (I - B_k Q_k B_k^T P_{k+1/k}^{-1})$ and

$$G_k = A_k^{-1} (I - B_k Q_k B_k^T P_{k+1/k}^{-1}). \quad (46)$$

In applications where difficulties exist with inverting $P_{k+1/k}$, it may be preferable to calculate

$$P_{k/k-1}^{-1} = P_{k/k}^{-1} - C_k^T R_k^{-1} C_k. \quad (47)$$

It is shown in [15] that the filter (37) - (38) and the smoother (39) can be written in the following Hamiltonian form

"Error is the discipline through which we advance." *William Ellery Channing*

$$\begin{bmatrix} \hat{x}_{k+1/N} \\ \lambda_{k/N} \end{bmatrix} = \begin{bmatrix} A_k & B_k Q_k B_k^T \\ -C_k^T R_k^{-1} C_k & A_k^T \end{bmatrix} \begin{bmatrix} \hat{x}_{k/N} \\ \lambda_{k+1/N} \end{bmatrix} + \begin{bmatrix} 0 \\ C_k^T R_k^{-1} z_k \end{bmatrix}, \quad (48)$$

$$(49)$$

where $\lambda_{k/N} \in \mathbb{R}^n$ is an auxiliary variable that proceeds backward in time k . The form (48) – (49) avoids potential numerical difficulties that may be associated with calculating $P_{k/k-1}^{-1}$.

To confirm the equivalence of (39) and (48) – (49), use the Bryson-Frazier formula [15]

$$\hat{x}_{k+1/N} = \hat{x}_{k+1/k+1} + P_{k+1/k} \lambda_{k+1/N}, \quad (50)$$

and (46) within (48) to obtain

$$\hat{x}_{k/N} = G_k \hat{x}_{k+1/k+1} + A_k^{-1} B_k Q_k B_k^T P_{k+1/k}^{-1} \hat{x}_{k+1/k}. \quad (51)$$

Employing (46) within (51) and rearranging leads to (39).

In time-invariant problems, steady state solutions for $P_{k/k}$ and $P_{k+1/k}$ can be used to precalculate the gain (40) before running the smoother. For example, the application of a time-invariant version of the Rauch-Tung-Striebel smoother for the restoration of blurred images is described in [14].

7.5.1.3 Performance

An expression for the smoother error covariance is developed below following the approach of [6], [13]. Define the smoother and filter error states as $\tilde{x}_{k/N} = x_k - \hat{x}_{k/N}$ and $\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}$, respectively. It is assumed that

$$E\{\tilde{x}_{k/k} \hat{x}_{k/k}^T\} = 0, \quad (52)$$

$$E\{\tilde{x}_{k+1/N} \hat{x}_{k+1/N}^T\} = 0, \quad (53)$$

$$E\{\tilde{x}_{k+1/N} \hat{x}_{k/N}^T\} = 0. \quad (54)$$

It is straightforward to show that (52) implies

$$E\{\hat{x}_{k+1/k} \hat{x}_{k+1/k}^T\} = E\{x_{k+1} x_{k+1}^T\} - P_{k+1/k}. \quad (55)$$

Denote $\Sigma_{k/N} = E\{\hat{x}_{k+1/N} \hat{x}_{k+1/N}^T\}$. The assumption (53) implies

$$E\{\hat{x}_{k+1/N} \hat{x}_{k+1/N}^T\} = E\{x_{k+1} x_{k+1}^T\} - \Sigma_{k+1/N}. \quad (56)$$

Subtracting x_k from both sides of (39) gives

$$\tilde{x}_{k/N} - G_k \hat{x}_{k+1/N} = \tilde{x}_{k/k} - G_k A_k \hat{x}_{k/k}. \quad (57)$$

By simplifying

“Great thinkers think inductively, that is, they create solutions and then seek out the problems that the solutions might solve; most companies think deductively, that is, defining a problem and then investigating different solutions.” *Joey Reiman*

$$E\{(\tilde{x}_{k/N} - G_k \hat{x}_{k+1/N})(\tilde{x}_{k/N} - G_k \hat{x}_{k+1/N})^T\} = E\{(\tilde{x}_{k/k} - G_k A_k \hat{x}_{k/k})(\tilde{x}_{k/k} - G_k A_k \hat{x}_{k/k})^T\} \quad (58)$$

and using (52), (54) – (56) yields

$$\Sigma_{k/N} = P_{k/k} - G_k(P_{k+1/k} - \Sigma_{k+1/N})G_k^T. \quad (59)$$

It can now be shown that the smoother performs better than the Kalman filter.

Lemma 4: Suppose that the sequence (59) is initialised with

$$\Sigma_{N+1/N} = P_{N+1/N}, \quad (60)$$

Then $\Sigma_{k/N} \leq P_{k/k}$ for $1 \leq k \leq N$.

Proof: The condition (60) implies $\Sigma_{N/N} = P_{N/N}$, which is the initial step for an induction argument. For the induction step, (59) is written as

$$\Sigma_{k/N} = P_{k/k-1} - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R)C_kP_{k/k-1} - G_k(P_{k+1/k} - \Sigma_{k+1/N})G_k^T \quad (61)$$

and thus $\Sigma_{k+1/N} \leq P_{k+1/k}$ implies $\Sigma_{k/N} \leq P_{k/k-1}$ and $\Sigma_{k/N} \leq P_{k/k}$. \square

7.5.2 The Fraser-Potter Smoother

Forward and backward estimates may be merged using the data fusion formula described in Lemma 7 of Chapter 6. A variation of the Fraser-Potter discrete-time fixed-interval smoother [4] derived by Monzingo [16] is advocated below.

In the first pass, a Kalman filter produces corrected state estimates $\hat{x}_{k/k}$ and corrected error covariances $P_{k/k}$ from the measurements. In the second pass, a Kalman filter is employed to calculate predicted “backward” state estimates $\xi_{k-1/k}$ and predicted “backward” error covariances $\Sigma_{k-1/k}$ from the time-reversed measurements. The smoothed estimate is given by [16]

$$\hat{x}_{k/N} = (P_{k/k}^{-1} + \Sigma_{k/k-1}^{-1})^{-1}(P_{k/k}^{-1}\hat{x}_{k/k} + \Sigma_{k/k-1}^{-1}\xi_{k/k-1}). \quad (62)$$

Alternatively, Kalman filters could be used to derive predicted quantities, $\hat{x}_{k/k-1}$ and $P_{k/k-1}$, from the measurements, and backward corrected quantities $\xi_{k/k}$ and $\Sigma_{k/k}$. Smoothed estimates may then be obtained from the linear combination

$$\hat{x}_{k/N} = (P_{k/k-1}^{-1} + \Sigma_{k/k}^{-1})^{-1}(P_{k/k-1}^{-1}\hat{x}_{k/k-1} + \Sigma_{k/k}^{-1}\xi_{k/k}). \quad (63)$$

It is observed that the fixed-point smoother (24), the fixed-lag smoother (32), maximum-likelihood smoother (39), the smoothed estimates (62) – (63) and the minimum-variance smoother (which is described subsequently) all use each measurement z_k once.

Note that Fraser and Potter’s original smoother solution [4] and Monzingo’s variation [16] are *ad hoc* and no claims are made about attaining a prescribed level of performance.

“No great discovery was ever made without a bold guess.” *Isaac Newton*

7.5.3 Minimum-Variance Smoothers

7.5.3.1 Optimal Unrealisable Solutions

Consider again the estimation problem depicted in Fig. 1 of Chapter 6, where w and v are now discrete-time inputs. As in continuous-time, it is desired to construct a solution \mathcal{H} is that produces output estimates \hat{y}_1 of a reference system $y_1 = \mathcal{G}_1 w$ from observations $z = y_2 + v$, where $y_2 = \mathcal{G}_2 w$. The objective is to minimise the energy of the output estimation error $e = y_1 - \hat{y}_1$.

The following discussion is perfunctory since it is a regathering of the results from Chapter 6. Recall that the output estimation error is generated by $e = \mathcal{R}_{ei} i$, where $\mathcal{R}_{ei} =$

$-\begin{bmatrix} \mathcal{H} & \mathcal{H}\mathcal{G}_2 - \mathcal{G}_1 \end{bmatrix}$ and $i = \begin{bmatrix} v \\ w \end{bmatrix}$. It has been shown previously that $\mathcal{R}_{ei} \mathcal{R}_{ei}^H = \mathcal{R}_{ei1} \mathcal{R}_{ei1}^H + \mathcal{R}_{ei2} \mathcal{R}_{ei2}^H$ where

$$\mathcal{R}_{ei2} = \mathcal{H}\Delta - \mathcal{G}_1 Q \mathcal{G}_2^H \Delta^{-H}, \quad (64)$$

in which $\Delta: \mathbb{R}^p \rightarrow \mathbb{R}^p$, is known as the Wiener-Hopf factor, which satisfies $\Delta \Delta^H = \mathcal{G}_1 Q \mathcal{G}_1^H + R$. From Lemma 8 of Chapter 6, the smoother solution $\mathcal{H} = \mathcal{G}_2 Q \mathcal{G}_2^H (\Delta \Delta^H)^{-1}$ achieves the best-possible performance, namely, it minimises $\|e\|_2 = \|\mathcal{R}_{ei} i\|_2 = \|\mathcal{R}_{ei} \mathcal{R}_{ei}^H\|_2^{1/2}$. For example, in output estimation problems $\mathcal{G}_1 = \mathcal{G}_2$ and the optimal smoother simplifies to $\mathcal{H}_{OE} = I - R(\Delta \Delta^H)^{-1}$. From Lemma 9 of Chapter 6, the (causal) filter solution $\mathcal{H}_{OE} = \{\mathcal{G}_2 Q \mathcal{G}_2^H \Delta^{-H} \Delta^{-1}\}_+ = \{\mathcal{G}_2 Q \mathcal{G}_2^H \Delta^{-H}\}_+ \Delta^{-1}$ achieves the best-possible filter performance, that is, it minimises $\|\{\tilde{y}_2 \tilde{y}_2^H\}_+\|_2 = \|\{\mathcal{R}_{ei} \mathcal{R}_{ei}^H\}_+\|_2$. The optimal smoother outperforms the optimal filter since $\|\{\mathcal{R}_{ei} \mathcal{R}_{ei}^H\}_+\|_2 \geq \|\mathcal{R}_{ei} \mathcal{R}_{ei}^H\|_2$. The above solutions are termed unrealisable because of the difficulty in obtaining Δ when \mathcal{G}_1 and \mathcal{G}_2 are time-varying systems. Realisable solutions that use an approximate Wiener-Hopf factor in place of Δ are presented below.

7.5.3.2 Non-causal Output Estimation

Suppose that the time-varying linear system \mathcal{G}_2 has the realisation (1) – (2). An approximate Wiener-Hopf factor $\hat{\Delta}: \mathbb{R}^p \rightarrow \mathbb{R}^p$ is introduced in [7], [13] and defined by

$$\begin{bmatrix} x_{k+1} \\ \delta_k \end{bmatrix} = \begin{bmatrix} A_k & K_k \Omega_k^{1/2} \\ C_k & \Omega_k^{1/2} \end{bmatrix} \begin{bmatrix} x_k \\ z_k \end{bmatrix}, \quad (65)$$

“Life is pretty simple. You do stuff. Most fails. Some works. You do more of what works. If it works big, others quickly copy it. Then you do something else. The trick is the doing something else.” *Leonardo da Vinci*

where $K_k = (A_k P_{k/k-1} C_k^T + B_k Q_k D_k^T) \Omega_k^{-1}$ is the predictor gain, $\Omega_k = C_k P_{k/k-1} C_k^T + D_k Q_k D_k^T + R_k$ and $P_{k/k-1}$ evolves from the Riccati difference equation $P_{k+1/k} = A_k P_{k/k-1} A_k^T - (A_k P_{k/k-1} C_k^T + B_k Q_k D_k^T)(C_k P_{k/k-1} C_k^T + D_k Q_k D_k^T + R_k)^{-1}(C_k P_{k/k-1} A_k^T + D_k Q_k B_k^T) + B_k Q_k B_k^T$. The inverse of the system (65), denoted by $\hat{\Delta}^{-1}$, is obtained using the Matrix Inversion Lemma

$$\begin{bmatrix} \hat{x}_{k+1/k} \\ \alpha_k \end{bmatrix} = \begin{bmatrix} A_k - K_k C_k & K_k \\ -\Omega_k^{-1/2} C_k & -\Omega_k^{-1/2} D_k \end{bmatrix} \begin{bmatrix} \hat{x}_{k/k-1} \\ z_k \end{bmatrix}. \quad (66)$$

The optimal output estimation smoother can be approximated as

$$\begin{aligned} \mathcal{H}_{OE} &= I - R(\hat{\Delta} \hat{\Delta}^H)^{-1} \\ &= I - R \hat{\Delta}^{-H} \hat{\Delta}^{-1}. \end{aligned} \quad (67)$$

A state-space realisation of (67) is given by (66),

$$\begin{bmatrix} \xi_{k-1} \\ \beta_k \end{bmatrix} = \begin{bmatrix} A_k^T - C_k^T K_k^T & C_k^T \Omega_k^{-1/2} \\ -K_k^T & \Omega_k^{-1/2} \end{bmatrix} \begin{bmatrix} \xi_k \\ \alpha_k \end{bmatrix}, \quad \xi_N = 0 \quad (68)$$

and

$$\hat{y}_{k/N} = z_k - R_k \beta_k. \quad (69)$$

Note that Lemma 1 is used to obtain the realisation (68) of $\hat{\Delta}^{-H} = (\hat{\Delta}^H)^{-1}$ from (66). A block diagram of this smoother is provided in Fig. 2. The states $\hat{x}_{k/k-1}$ within (66) are immediately recognisable as belonging to the one-step-ahead predictor. Thus, the optimum realisable solution involves a cascade of familiar building blocks, namely, a Kalman predictor and its adjoint.

Procedure 1. The above output estimation smoother can be implemented via the following three-step procedure.

- Step 1. Operate $\hat{\Delta}^{-1}$ on z_k using (66) to obtain α_k .
- Step 2. In lieu of the adjoint system (68), operate (66) on the time-reversed transpose of α_k . Then take the time-reversed transpose of the result to obtain β_k .
- Step 3. Calculate the smoothed output estimate from (69).

It is shown below that $\hat{y}_{k/N}$ is an unbiased estimate of y_k .

“When I am working on a problem, I never think about beauty but when I have finished, if the solution is not beautiful, I know it is wrong.” *Richard Buckminster Fuller*

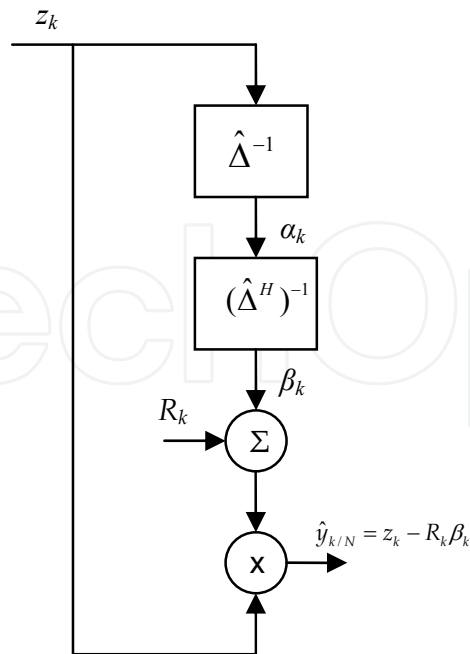


Figure 2. Block diagram of the output estimation smoother

Lemma 5 $E\{\hat{y}_{k/N}\} = E\{y_k\}$.

Proof: Denote the one-step-ahead prediction error by $\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1}$. The output of (66) may be written as $\alpha_k = \Omega_k^{-1/2} C_k \tilde{x}_{k/k-1} + \Omega_k^{-1/2} v_k$ and therefore

$$E\{\alpha_k\} = \Omega_k^{-1/2} C_k E\{\tilde{x}_{k/k-1}\} + \Omega_k^{-1/2} E\{v_k\}. \quad (70)$$

The first term on the right-hand-side of (70) is zero since it pertains to the prediction error of the Kalman filter. The second term is zero since it is assumed that $E\{v_k\} = 0$. Thus $E\{\alpha_k\} = 0$. Since the recursion (68) is initialized with $\zeta_N = 0$, it follows that $E\{\zeta_k\} = 0$, which implies $E\{\zeta_k\} = -K_k E\{\zeta_k\} + \Omega_k^{-1/2} E\{\alpha_k\} = 0$. Thus, from (69), $E\{\hat{y}_{k/N}\} = E\{z_k\} = E\{y_k\}$, since it is assumed that $E\{v_k\} = 0$. \square

7.5.3.3 Causal Output Estimation

The minimum-variance (Kalman) filter is obtained by taking the causal part of the optimum minimum-variance smoother (67)

$$\{\mathcal{H}_{OE}\}_+ = I - R_k \{\hat{\Delta}^{-H}\}_+ \hat{\Delta}^{-1} \quad (71)$$

“The practical success of an idea, irrespective of its inherent merit, is dependent on the attitude of the contemporaries. If timely it is quickly adopted; if not, it is apt to fare like a sprout lured out of the ground by warm sunshine, only to be injured and retarded in its growth by the succeeding frost.” *Nicola Tesla*

$$= I - R_k \Omega_k^{-1/2} \hat{\Delta}^{-1}.$$

To confirm this linkage between the smoother and filter, denote $L_k = (C_k P_{k/k-1} C_k^T + D_k Q_k D_k^T) \Omega_k^{-1}$ and use (71) to obtain

$$\begin{aligned} \hat{y}_{k/k} &= z_k - R_k \Omega_k^{-1/2} \alpha_k \\ &= R_k \Omega_k^{-1} C_k x_k + (I - R_k \Omega_k^{-1/2}) z_k \\ &= (C_k - L_k C_k) x_k + L_k z_k, \end{aligned} \quad (72)$$

which is identical to (34) of Chapter 4.

7.5.3.4 Input Estimation

As discussed in Chapter 6, the optimal realisable smoother for input estimation is

$$\mathcal{H}_{IE} = Q \mathcal{G}_2^H \hat{\Delta}^{-H} \hat{\Delta}^{-1}. \quad (73)$$

The development of a state-space realisation for $\hat{w}_{k/N} = Q \mathcal{G}_2^H \hat{\Delta}^{-H} \alpha_k$ makes use of the formula for the cascade of two systems described in Chapter 6. The smoothed input estimate is realised by

$$\begin{bmatrix} \xi_{k-1} \\ \gamma_{k-1} \\ \hat{w}_{k-1/N} \end{bmatrix} = \begin{bmatrix} A_k^T - C_k^T K_k & 0 & C_k^T \Omega_k^{-1/2} \\ C_k^T K_k^T & A_k^T & -C_k^T \Omega_k^{-1/2} \\ -Q_k D_k^T K_k^T & -Q_k B_k^T & Q_k D_k^T \Omega_k^{-1/2} \end{bmatrix} \begin{bmatrix} \xi_k \\ \gamma_k \\ \alpha_k \end{bmatrix}, \quad (74)$$

in which $\gamma_k \in \mathbb{R}^n$ is an auxiliary state.

Procedure 2. The above input estimator can be implemented via the following three steps.

- Step 1. Operate $\hat{\Delta}^{-1}$ on the measurements z_k using (66) to obtain α_k .
- Step 2. Operate the adjoint of (74) on the time-reversed transpose of α_k . Then take the time-reversed transpose of the result.

7.5.3.5 State Estimation

Smoothed state estimates can be obtained by defining the reference system $\mathcal{G}_1 = I$ which yields

$$\begin{aligned} \hat{x}_{k+1/N} &= A_k \hat{x}_{k/N} + B_k \hat{w}_{k/N} \\ &= A_k \hat{x}_{k/N} + B_k Q \mathcal{G}_2^H \hat{\Delta}^{-H} \alpha_k. \end{aligned} \quad (75)$$

"Doubt is the father of invention." Galileo Galilei

Thus, the minimum-variance smoother for state estimation is given by (66) and (74) – (75). As remarked in Chapter 6, some numerical model order reduction may be required. In the special case of C_k being of rank n and $D_k = 0$, state estimates can be calculated from (69) and

$$\hat{x}_{k/N} = C_k^\# \hat{y}_{k/N} . \quad (76)$$

where $C_k^\# = (C_k^T C_k)^{-1} C_k^T$ is the Moore-Penrose pseudo-inverse.

7.5.3.6 Performance

The characterisation of smoother performance requires the following additional notation. Let $\gamma = \mathcal{G}_0 w$ denote the output of the linear time-varying system having the realisation

$$x_{k+1} = A_k x_k + w_k , \quad (77)$$

$$\gamma_k = x_k , \quad (78)$$

where $A_k \in \mathbb{R}^{n \times n}$. By inspection of (77) – (78), the output of the inverse system $w = \mathcal{G}_0^{-1} y$ is given by

$$w_k = \gamma_{k+1} - A_k \gamma_k . \quad (79)$$

Similarly, let $\varepsilon = \mathcal{G}_0^H u$ denote the output of the adjoint system \mathcal{G}_0^H , which from Lemma 1 has the realisation

$$\zeta_{k-1} = A_k^T \zeta_k - u_k , \quad (80)$$

$$\varepsilon_k = -\zeta_k . \quad (81)$$

It follows that the output of the inverse system $u = \mathcal{G}_0^{-H} \varepsilon$ is given by

$$u_k = \varepsilon_{k-1} - A_k^T \varepsilon_k . \quad (82)$$

The exact Wiener-Hopf factor may now be written as

$$\Delta \Delta^H = C_k \mathcal{G}_0 B_k Q_k B_k^T \mathcal{G}_0^H C_k^T + R_k . \quad (83)$$

The subsequent lemma, which relates the exact and approximate Wiener-Hopf factors, requires the identity

$$P_k - A_k P_k A_k^T = A_k P_k \mathcal{G}_0^{-H} + \mathcal{G}_0^{-1} P_k A_k^T + \mathcal{G}_0^{-1} P_k \mathcal{G}_0^{-H} , \quad (84)$$

in which P_k is an arbitrary matrix of appropriate dimensions. A verification of (84) is requested in the problems.

“The theory of our modern technique shows that nothing is as practical as the theory.” *Julius Robert Oppenheimer*

Lemma 6 [7]: In respect of the signal model (1) – (2) with $D_k = 0$, $E\{w_k\} = E\{v_k\} = 0$, $E\{w_j w_k^T\} = Q_k \delta_{jk}$, $E\{v_j v_k^T\} = R_k \delta_{jk}$, $E\{w_k v_k^T\} = 0$ and the quantities defined above,

$$\hat{\Delta} \hat{\Delta}^H = \Delta \Delta^H + C_k \mathcal{G}_0 (P_{k/k-1} - P_{k+1/k}) \mathcal{G}_0^H C_k^T. \quad (85)$$

Proof: The approximate Wiener-Hopf factor may be written as $\hat{\Delta} = C_k \mathcal{G}_0 K_k \Omega_k^{1/2} + \Omega_k^{1/2}$. Using $P_{k/k-1} - A_k P_{k/k-1} A_k^T = -A_k P_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + R_k)^{-1} C_k P_{k/k-1} A_k^T + B_k Q_k B_k^T + P_{k/k-1} - P_{k+1/k}$ within (84) and simplifying leads to (85). \square

It can be seen from (85) that the approximate Wiener-Hopf factor approaches the exact Wiener-Hopf factor whenever the estimation problem is locally stationary, that is, when the model and noise parameters vary sufficiently slowly so that $P_{k/k-1} \approx P_{k+1/k}$. Under these conditions, the smoother (69) achieves the best-possible performance, as is shown below.

Lemma 7 [7]: The smoother (67) satisfies

$$\mathcal{R}_{ei2} = R_k [(\Delta \Delta^H)^{-1} - (\Delta \Delta^H - C_k \mathcal{G}_0 (P_{k/k-1} - P_{k+1/k}) \mathcal{G}_0^H C_k^T)^{-1}] \Delta, \quad (86)$$

Proof: Substituting (67) into (64) yields

$$\mathcal{R}_{ei2} = R_k [(\Delta \Delta^H)^{-1} - (\hat{\Delta} \hat{\Delta}^H)^{-1}] \Delta. \quad (87)$$

The result (86) is now immediate from (85) and (87). \square

Some conditions under which $P_{k+1/k}$ asymptotically approaches $P_{k/k-1}$ and the smoother (67) attaining optimal performance are set out below.

Lemma 8 [8]: Suppose

- (i) for a $t > 0$ that there exist solutions $P_t \geq P_{t+1} \geq 0$ of the Riccati difference equation
- (ii) $P_{k+1/k} = A_k P_{k/k-1} A_k^T - A_k P_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + R_k)^{-1} C_k P_{k/k-1} A_k^T + B_k Q_k B_k^T$; and
- (iii) $\begin{bmatrix} B_{t+k-1} Q_{t+k-1} B_{t+k-1}^T & A_{t+k-1}^T \\ A_{t+k-1} & -C_{t+k-1}^T R_{t+k-1}^{-1} C_{t+k-1} \end{bmatrix} \geq \begin{bmatrix} B_{t+k} Q_{t+k} B_{t+k}^T & A_{t+k}^T \\ A_{t+k} & -C_{t+k}^T R_{t+k}^{-1} C_{t+k} \end{bmatrix}$ for all $k \geq 0$.

Then the smoother (67) achieves

$$\lim_{t \rightarrow \infty} \|\mathcal{R}_{ei2} \mathcal{R}_{ei2}^H\|_2 = 0. \quad (88)$$

Proof: Conditions i) and ii) together with Theorem 1 imply $P_{k/k-1} \geq P_{k+1/k}$ for all $k \geq 0$ and $P_{k/k-1} = 0$. The claim (88) is now immediate from Theorem 2. \square

An example that illustrates the performance benefit of the minimum-variance smoother (67) is described below.

"Whoever, in the pursuit of science, seeks after immediate practical utility may rest assured that he seeks in vain." Hermann Ludwig Ferdinand von Helmholtz

Example 2 [9]. The nominal drift rate of high quality inertial navigation systems is around one nautical mile per hour, which corresponds to position errors of about 617 m over a twenty minute period. Thus, inertial navigation systems alone cannot be used to control underground mining equipment. An approach which has been found to be successful in underground mines is called dead reckoning, where the Euler angles, θ_k , φ_k and ϕ_k , reported by an inertial navigation system are combined with external odometer measurements, d_k . The dead reckoning position estimates in the x-y-z plane are calculated as

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} + (d_k - d_{k-1}) \begin{bmatrix} \sin(\theta_k) \\ \sin(\varphi_k) \\ \sin(\phi_k) \end{bmatrix}. \quad (89)$$

A filter or a smoother may then be employed to improve the noisy position estimates

calculated from (89). Euler angles were generated using $\begin{bmatrix} \theta_{k+1} \\ \varphi_{k+1} \\ \phi_{k+1} \end{bmatrix} = A \begin{bmatrix} \theta_k \\ \varphi_k \\ \phi_k \end{bmatrix} + \begin{bmatrix} w_k^{(1)} \\ w_k^{(2)} \\ w_k^{(3)} \end{bmatrix}$, with

$A = \begin{bmatrix} 0.95 & 0 & 0 \\ 0 & 0.95 & 0 \\ 0 & 0 & 0.95 \end{bmatrix}$ and $w_k^{(i)} \sim \mathcal{N}(0, 0.01)$, $i = 1 \dots 3$. Simulations were conducted with

1000 realisations of Gaussian measurement noise added to position estimates calculated from (89). The mean-square error exhibited by the minimum-variance filter and smoother operating on the noisy dead reckoning estimates are shown in Fig. 3. It can be seen that filtering the noisy dead reckoning positions can provide a significant mean-square-error improvement. The figure also demonstrates that the smoother can offer a few dB of further improvement at mid-range signal-to-noise ratios.

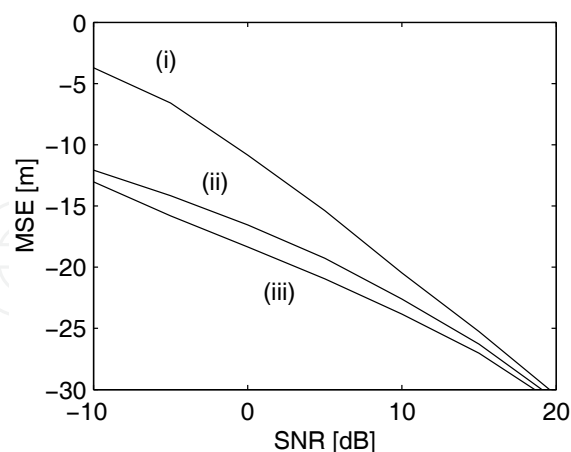


Figure 3. Mean-square-error of the position estimate versus input signal to noise ratio for Example 2: (i) noisy dead reckoning data, (ii) Kalman filter, and (iii) minimum-variance smoother (69).

"I do not think that the wireless waves that I have discovered will have any practical application."
Heinrich Rudolf Hertz

7.6 Performance Comparison

It has been demonstrated by the previous examples that the optimal fixed-interval smoother provides a performance improvement over the maximum-likelihood smoother. The remaining example of this section compares the behaviour of the fixed-lag and the optimum fixed-interval smoother.

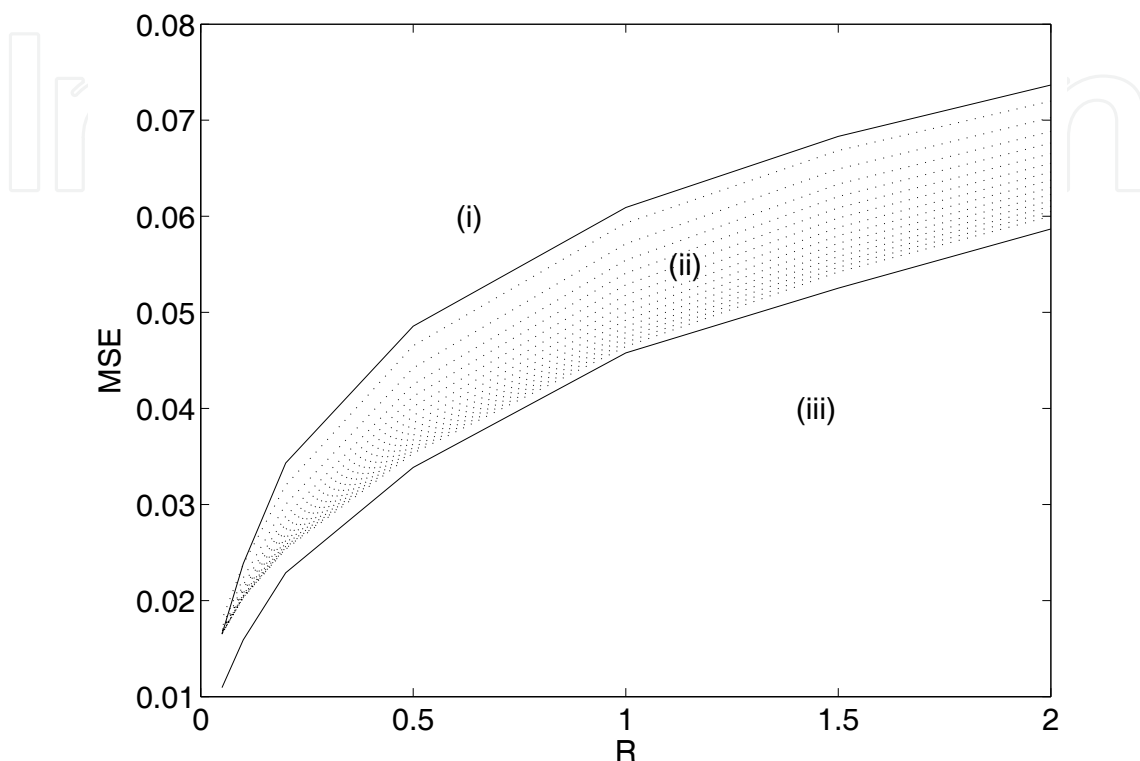


Figure 4. Mean-square-error versus measurement noise covariance for Example 3: (i) Kalman filter, (ii) fixed-lag smoothers, and (iii) optimal minimum-variance smoother (67).

Example 3. Simulations were conducted for a first-order output estimation problem, in which $A = 0.95$, $B = 1$, $C = 0.1$, $Q = 1$, $R = \{0.01, 0.02, 0.5, 1, 1.5, 2\}$ and $N = 20,000$. The mean-square-errors exhibited by the Kalman filter and the optimum fixed-interval smoother (69) are indicated by the top and bottom solid lines of Fig. 4, respectively. Fourteen fixed-lag smoother output error covariances, $CP_{k+1/k}^{(i+1,i+1)}C^T$, $i = 2 \dots 15$, were calculated from (35) – (36) and are indicated by the dotted lines of Fig. 4. The figure illustrates that the fixed-lag smoother mean-square errors are bounded above and below by those of the Kalman filter and optimal fixed-interval smoother, respectively. Thus, an option for asymptotically attaining optimal performance is to employ Moore's reduced-order fixed-lag solution [1] – [3] with a sufficiently long lag.

"You see, wire telegraph is a kind of a very, very long cat. You pull his tail in New York and his head is meowing in Los Angeles. Do you understand this? And radio operates exactly the same way: you send signals here, they receive them there. The only difference is that there is no cat." *Albert Einstein*

	ASSUMPTIONS	MAIN RESULTS
Signals and system	$E\{w_k\} = E\{v_k\} = 0$. $E\{w_k w_k^T\} = Q_k > 0$ and $E\{v_k v_k^T\} = R_k > 0$ are known. $A_k, B_k, C_{1,k}$ and $C_{2,k}$ are known.	$x_{k+1} = A_k x_k + B_k w_k$ $y_{2,k} = C_{2,k} x_k$ $z_k = y_{2,k} + v_k$ $y_{1,k} = C_{1,k} x_k$
RTS smoother	Assumes that the filtered and smoothed states are normally distributed. $\hat{x}_{k+1/k}$ previously calculated by Kalman filter.	$\hat{x}_{k/N} = \hat{x}_{k/k} + G_k (\hat{x}_{k+1/N} - \hat{x}_{k+1/k})$ $\hat{y}_{1,k/N} = C_{1,k} \hat{x}_{k/N}$
FP smoother	$\hat{x}_{k/k}$ and $\xi_{k/k-1}$ previously calculated by forward and backward Kalman filters.	$\hat{x}_{k/N} = (P_{k/k}^{-1} + \Sigma_{k/k-1}^{-1})^{-1} (P_{k/k}^{-1} \hat{x}_{k/k} + \Sigma_{k/k-1}^{-1} \xi_{k/k-1})$ $\hat{y}_{1,k/N} = C_{1,k} \hat{x}_{k/N}$
Optimal minimum-variance smoother		$\begin{bmatrix} x_{k+1/k} \\ \alpha_k \end{bmatrix} = \begin{bmatrix} A_k - K_k C_{2,k} & K_k \\ -\Omega_k^{-1/2} C_{2,k} & \Omega_k^{-1/2} \end{bmatrix} \begin{bmatrix} x_{k/k-1} \\ z_k \end{bmatrix}$ $\begin{bmatrix} \xi_{k-1} \\ \beta_k \end{bmatrix} = \begin{bmatrix} A_k^T - C_{2,k}^T K_k^T & C_{2,k}^T \Omega_k^{-1/2} \\ -K_k^T & \Omega_k^{-1/2} \end{bmatrix} \begin{bmatrix} \xi_k \\ \alpha_k \end{bmatrix}$ $\hat{y}_{2,k/N} = z_k - R_k \beta_k$ $\hat{w}_{k/N} = Q \mathcal{G}_2^H \hat{\Delta}^{-H} \alpha_k$ $\hat{x}_{k+1/N} = A_k \hat{x}_{k/N} + B_k \hat{w}_{k/N}$ $\hat{y}_{1,k/N} = C_{1,k} \hat{x}_{k/N}$

Table 1. Main results for discrete-time fixed-interval smoothing.

7.7 Conclusion

Solutions to the fixed-point and fixed-lag smoothing problems can be found by applying the standard Kalman filter recursions to augmented systems. Where possible, it is shown that the smoother error covariances are less than the filter error covariance, namely, the fixed-point and fixed-lag smoothers provide a performance improvement over the filter.

Table 1 summarises three common fixed-interval smoothers that operate on measurements $z_k = y_{2,k} + v_k$ of a system \mathcal{G}_2 realised by $x_{k+1} = A_k x_k + B_k w_k$ and $y_{2,k} = C_{2,k} x_k$. Monzingo modified the Fraser-Potter smoother solution so that each measurement is only used once. Rauch, Tung and Striebel employed the maximum-likelihood method to derive their fixed-interval

“To invent, you need a good imagination and a pile of junk.” Thomas Alva Edison

smoother in which $G_k = P_{k/k} A_{k+1}^T P_{k+1/k}^{-1}$ is a gain matrix. Although this is not a minimum-mean-square-error solution, it outperforms the Kalman filter and can provide close to optimal performance whenever the underlying assumptions are reasonable.

The minimum-variance smoothers are state-space generalisations of the optimal noncausal Wiener solutions. They make use of the inverse of the approximate Wiener-Hopf factor $\hat{\Delta}^{-1}$ and its adjoint $\hat{\Delta}^{-H}$. These smoothers achieve the best-possible performance, however, they are not minimum-order solutions. Consequently, any performance benefits need to be reconciled against the increased complexity.

7.8 Problems

Problem 1.

- (i) Simplify the fixed-lag smoother

$$\begin{bmatrix} \hat{x}_{k+1/k} \\ \hat{x}_{k/k} \\ \hat{x}_{k-1/k} \\ \vdots \\ \hat{x}_{k-N+1/k} \end{bmatrix} = \begin{bmatrix} A_k & 0 & \cdots & 0 \\ I_n & 0 & & 0 \\ 0 & I_n & & \vdots \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & I_n & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{k/k-1} \\ \hat{x}_{k-1/k-1} \\ \hat{x}_{k-2/k-1} \\ \vdots \\ \hat{x}_{k-N/k-1} \end{bmatrix} + \begin{bmatrix} K_{0,k} \\ K_{1,k} \\ K_{2,k} \\ \vdots \\ K_{N,k} \end{bmatrix} (z_k - C_k \hat{x}_{k/k-1})$$

to obtain an expression for the components of the smoothed state.

- (ii) Derive expressions for the two predicted error covariance submatrices of interest.

Problem 2.

- (i) With the quantities defined in Section 4 and the assumptions $\hat{x}_{k+1/N} \sim \mathcal{N}(A_k \hat{x}_{k/N}, B_k Q_k B_k^T)$, $\hat{x}_{k/N} \sim \mathcal{N}(\hat{x}_{k/k}, P_{k/k})$, use the maximum-likelihood method to derive

$$\hat{x}_{k/N} = (I + P_{k/k} A_k^T (B_k Q_k B_k^T)^{-1} A_k)^{-1} (\hat{x}_{k/N} + P_{k/k} A_k^T B_k Q_k B_k^T)^{-1} \hat{x}_{k/N})^T.$$

- (i) Use the Matrix Inversion Lemma to obtain Rauch, Tung and Striebel's smoother

$$\hat{x}_{k/N} = \hat{x}_{k/k} + G_k (\hat{x}_{k+1/N} - \hat{x}_{k+1/k}).$$

- (ii) Employ the additional assumptions $E\{\tilde{x}_{k/k} \hat{x}_{k/k}^T\} = 0$, $E\{\tilde{x}_{k+1/N} \hat{x}_{k+1/N}^T\}$ and $E\{\tilde{x}_{k+1/N} \hat{x}_{k/N}^T\}$ to show that $E\{\hat{x}_{k+1/k} \hat{x}_{k+1/k}^T\} = E\{x_{k+1} x_{k+1}^T\} - P_{k+1/k}$, $E\{\hat{x}_{k+1/N} \hat{x}_{k+1/N}^T\} = E\{x_{k+1} x_{k+1}^T\} - \Sigma_{k+1/N}$ and $\Sigma_{k/N} = P_{k/k} - G_k (P_{k+1/k} - \Sigma_{k+1/N}) G_k^T$.

"My invention, (the motion picture camera), can be exploited for a certain time as a scientific curiosity, but apart from that it has no commercial value whatsoever." August Marie Louis Nicolas Lumiere

(iii) Use $G_k = A_k^{-1}(I - B_k Q_k B_k^T P_{k+1/k}^{-1})$ and $\hat{x}_{k+1/N} = \hat{x}_{k+1/k+1} + P_{k+1/k} \lambda_{k+1/N}$ to confirm that the smoothed estimate within

$$\begin{bmatrix} \hat{x}_{k+1/N} \\ \lambda_{k/N} \end{bmatrix} = \begin{bmatrix} A_k & B_k Q_k B_k^T \\ -C_k^T R_k^{-1} C_k & A_k^T \end{bmatrix} \begin{bmatrix} \hat{x}_{k+1/N} \\ \lambda_{k+1/N} \end{bmatrix} + \begin{bmatrix} 0 \\ C_k^T R_k^{-1} z_k \end{bmatrix}$$

is equivalent to Rauch, Tung and Striebel's maximum-likelihood solution.

Problem 3. Let $a = \mathcal{G}_0 w$ denote the output of linear time-varying system having the realisation $x_{k+1} = A_k x_k + w_k$, $y_k = x_k$. Verify that $P_k - A_k P_k A_k^T = A_k P_k \mathcal{G}_0^{-H} + \mathcal{G}_0^{-1} P_k A_k^T + \mathcal{G}_0^{-1} P_k \mathcal{G}_0^{-H}$.

Problem 4. For the model (1) - (2). assume that $D_k = 0$, $E\{w_k\} = E\{v_k\} = 0$, $E\{w_k w_k^T\} = Q_k \delta_{jk}$, $E\{v_j v_k^T\} = R_k \delta_{jk}$ and $E\{w_k v_k^T\} = 0$. Use the result of Problem 3 to show that $\hat{\Delta} \hat{\Delta}^H = \Delta \Delta^H - C_k \mathcal{G}_0 (P_{k/k-1} - P_{k+1/k}) \mathcal{G}_0^H C_k^T$.

Problem 5. Under the assumptions of Problem 4, obtain an expression relating $\hat{\Delta} \hat{\Delta}^H$ and $\Delta \Delta^H$ for the case where $D_k \neq 0$.

Problem 6. From $\mathcal{R}_{ei} = -[\mathcal{H} \quad \mathcal{H} \mathcal{G}_2 - \mathcal{G}_1]$, $\mathcal{R}_{ei} \mathcal{R}_{ei}^H = \mathcal{R}_{ei1} \mathcal{R}_{ei1}^H + \mathcal{R}_{ei2} \mathcal{R}_{ei2}^H$ and $\mathcal{R}_{ei2} = \mathcal{H} \Delta - \mathcal{G}_1 Q \mathcal{G}_2^H \Delta^{-H}$, obtain the optimal realisable smoother solutions for output estimation, input estimation and state estimation problems.

7.9 Glossary

$p(x_k)$	Probability density function of a discrete random variable x_k .
$x_k \sim \mathcal{N}(\mu, R_{xx})$	The random variable x_k has a normal distribution with mean μ and covariance R_{xx} .
$\Omega_{k/k}$	Error covariance of the fixed-point smoother.
$P_{k+1/k}^{(i,i)}$	Error covariance of the fixed-lag smoother.
$\hat{w}_{k/N}, \hat{x}_{k/N}, \hat{y}_{k/N}$	Estimates of w_k , x_k and y_k at time k , given data z_k over an interval $k \in [0, N]$.
G_k	Gain of the smoother developed by Rauch, Tung and Striebel.
α_k	Output of $\hat{\Delta}^{-1}$, the inverse of the approximate Wiener-Hopf factor.
β_k	Output of $\hat{\Delta}^{-H}$, the adjoint of the inverse of the approximate Wiener-Hopf factor.
$C_k^\#$	Moore-Penrose pseudoinverse of C_k .

"The cinema is little more than a fad. It's canned drama. What audiences really want to see is flesh and blood on the stage." *Charles Spencer (Charlie) Chaplin*

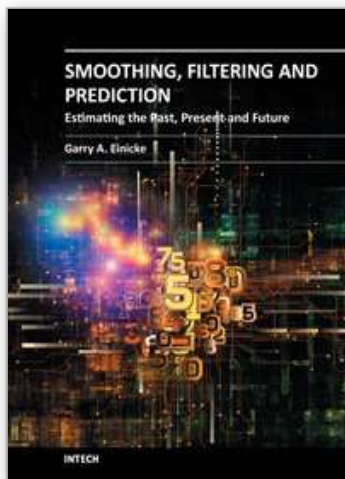
\mathcal{R}_{ei}

A system (or map) that operates on the problem inputs $i = \begin{bmatrix} v \\ w \end{bmatrix}$ to produce an estimation error e . It is convenient to make use of the factorisation $\mathcal{R}_{ei}\mathcal{R}_{ei}^H = \mathcal{R}_{ei1}\mathcal{R}_{ei1}^H + \mathcal{R}_{ei2}\mathcal{R}_{ei2}^H$, where $\mathcal{R}_{ei2}\mathcal{R}_{ei2}^H$ includes the filter/smoothing solution and $\mathcal{R}_{ei1}\mathcal{R}_{ei1}^H$ is a lower performance bound.

7.10 References

- [1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1979.
- [2] J. B. Moore, "Fixed-Lag Smoothing Results for Linear Dynamical Systems", *A.T.R.*, vol. 7, no. 2, pp. 16 – 21, 1973.
- [3] J. B. Moore, "Discrete-Time Fixed-Lag Smoothing Algorithms", *Automatica*, vol. 9, pp. 163 – 173, 1973.
- [4] D. C. Fraser and J. E. Potter, "The Optimum Linear Smoother as a Combination of Two Optimum Linear Filters", *IEEE Transactions on Automatic Control*, vol. AC-14, no. 4, pp. 387 – 390, Aug., 1969.
- [5] H. E. Rauch, "Solutions to the Linear Smoothing Problem", *IEEE Transactions on Automatic Control*, vol. 8, no. 4, pp. 371 – 372, Oct. 1963.
- [6] H. E. Rauch, F. Tung and C. T. Striebel, "Maximum Likelihood Estimates of Linear Dynamic Systems", *AIAA Journal*, vol. 3, no. 8, pp. 1445 – 1450, Aug., 1965.
- [7] G. A. Einicke, "Optimal and Robust Noncausal Filter Formulations", *IEEE Transactions on Signal Process.*, vol. 54, no. 3, pp. 1069 – 1077, Mar. 2006.
- [8] G. A. Einicke, "Asymptotic Optimality of the Minimum Variance Fixed-Interval Smoother", *IEEE Transactions on Signal Process.*, vol. 55, no. 4, pp. 1543 – 1547, Apr. 2007.
- [9] G. A. Einicke, J. C. Ralston, C. O. Hargrave, D. C. Reid and D. W. Hainsworth, "Longwall Mining Automation, An Application of Minimum-Variance Smoothing", *IEEE Control Systems Magazine*, vol. 28, no. 6, pp. 28 – 37, Dec. 2008.
- [10] H. K. Wimmer, "Monotonicity and Maximality of Solutions of Discrete-Time Algebraic Riccati Equations", *Journal of Mathematics, Systems, Estimation and Control*, vol. 2, no. 2, pp. 219 – 235, 1992.
- [11] H. K. Wimmer and M. Pavon, "A comparison theorem for matrix Riccati difference equations", *Systems & Control Letters*, vol. 19, pp. 233 – 239, 1992.
- [12] L. E. Zachrisson, "On Optimal Smoothing of Continuous Time Kalman Processes", *Information Sciences*, vol. 1, pp. 143 – 172, 1969.
- [13] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill Book Company, New York, 1971.
- [14] A. Asif, "Fast Rauch-Tung-Striebel Smoother-Base Image Restoration for Noncausal Images", *IEEE Signal Processing Letters*, vol. 11, no. 3, pp. 371 – 374, Mar. 2004.
- [15] T. Kailath, A. H. Sayed and B. Hassibi, *Linear Estimation*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2000.
- [16] R. A. Monzingo, "Discrete Optimal Linear Smoothing for Systems with Uncertain Observations", *IEEE Transactions on Information Theory*, vol. 21, no. 3, pp. 271 – 275, May 1975.

"Who the hell wants to hear actors talk?" Harry Morris Warner



Smoothing, Filtering and Prediction - Estimating The Past, Present and Future

Edited by

ISBN 978-953-307-752-9

Hard cover, 276 pages

Publisher InTech

Published online 24, February, 2012

Published in print edition February, 2012

This book describes the classical smoothing, filtering and prediction techniques together with some more recently developed embellishments for improving performance within applications. It aims to present the subject in an accessible way, so that it can serve as a practical guide for undergraduates and newcomers to the field. The material is organised as a ten-lecture course. The foundations are laid in Chapters 1 and 2, which explain minimum-mean-square-error solution construction and asymptotic behaviour. Chapters 3 and 4 introduce continuous-time and discrete-time minimum-variance filtering. Generalisations for missing data, deterministic inputs, correlated noises, direct feedthrough terms, output estimation and equalisation are described. Chapter 5 simplifies the minimum-variance filtering results for steady-state problems. Observability, Riccati equation solution convergence, asymptotic stability and Wiener filter equivalence are discussed. Chapters 6 and 7 cover the subject of continuous-time and discrete-time smoothing. The main fixed-lag, fixed-point and fixed-interval smoother results are derived. It is shown that the minimum-variance fixed-interval smoother attains the best performance. Chapter 8 attends to parameter estimation. As the above-mentioned approaches all rely on knowledge of the underlying model parameters, maximum-likelihood techniques within expectation-maximisation algorithms for joint state and parameter estimation are described. Chapter 9 is concerned with robust techniques that accommodate uncertainties within problem specifications. An extra term within Riccati equations enables designers to trade-off average error and peak error performance. Chapter 10 rounds off the course by applying the afore-mentioned linear techniques to nonlinear estimation problems. It is demonstrated that step-wise linearisations can be used within predictors, filters and smoothers, albeit by forsaking optimal performance guarantees.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Garry Einicke (2012). Discrete-Time Smoothing, Smoothing, Filtering and Prediction - Estimating The Past, Present and Future, (Ed.), ISBN: 978-953-307-752-9, InTech, Available from:

<http://www.intechopen.com/books/smoothing-filtering-and-prediction-estimating-the-past-present-and-future/discrete-time-smoothing>

INTech
open science | open minds

InTech Europe

University Campus STeP Ri

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai

www.intechopen.com

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen

IntechOpen

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen